

CALIBRATED ULTRA FAST IMAGE SIMULATIONS FOR THE DARK ENERGY SURVEY

CLAUDIO BRUDERER^{1†}, CHIHWAY CHANG¹, ALEXANDRE REFREGIER¹, ADAM AMARA¹, JOEL BERGE^{1,2}, AND LUKAS GAMPER¹¹ Institute for Astronomy, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zürich, Switzerland and² ONERA - The French Aerospace Lab, 29 avenue de la Division Leclerc, 92320 Châtillon, France

Draft version April 14, 2015

ABSTRACT

Weak lensing by large-scale structure is a powerful technique to probe the dark components of the universe. To understand the measurement process of weak lensing and the associated systematic effects, image simulations are becoming increasingly important. For this purpose we present a first implementation of the *Monte Carlo Control Loops* (MCCL; Refregier & Amara 2014), a coherent framework for studying systematic effects in weak lensing. It allows us to model and calibrate the shear measurement process using image simulations from the Ultra Fast Image Generator (UFIG; Bergé et al. 2013). We apply this framework to a subset of the data taken during the Science Verification period (SV) of the Dark Energy Survey (DES). We calibrate the UFIG simulations to be statistically consistent with DES images. We then perform tolerance analyses by perturbing the simulation parameters and study their impact on the shear measurement at the one-point level. This allows us to determine the relative importance of different input parameters to the simulations. For spatially constant systematic errors and six simulation parameters, the calibration of the simulation reaches the weak lensing precision needed for the DES SV survey area. Furthermore, we find a sensitivity of the shear measurement to the intrinsic ellipticity distribution, and an interplay between the magnitude-size and the pixel value diagnostics in constraining the noise model. This work is the first application of the MCCL framework to data and shows how it can be used to methodically study the impact of systematics on the cosmic shear measurement.

Subject headings: Gravitational lensing: weak — methods: numerical — methods: statistical — surveys

1. INTRODUCTION

Within the last decades our picture of the Universe changed dramatically with the discovery of its accelerating expansion attributed to a mysterious dark energy. Together with dark matter they make up the dark sector of the Universe. The introduction of dark energy has led to the establishment of the Λ CDM-model as the current cosmological standard model. The model agrees well with observations from different cosmological probes (e.g. Hinshaw et al. 2013; Planck Collaboration et al. 2014). Nonetheless, understanding the nature of the dark sector is one of cosmology's most pressing challenges.

Weak gravitational lensing (for reviews see Refregier 2003; Hoekstra & Jain 2008), a distortion effect of galaxy shapes due to interloping structures along the line-of-sight, has a large potential to shed light onto the mystery of the dark sector (Albrecht et al. 2006). It is a purely gravitational effect, and thus reacts in the same way to dark and baryonic matter. However, the induced distortions on galaxy shapes are very weak ($\sim 1\%$). Therefore, galaxy shapes need to be measured to a very high precision for weak lensing to unleash its full potential as a cosmological probe (e.g. Huterer et al. 2006; Amara & Réfrégier 2008, henceforth AR08).

Many shape-measurement algorithms have been developed over the past two decades (for an overview see Zuntz et al. 2013, and references therein). Image simulation plays an important role for calibrating and validating those methods. To test the performance of vari-

ous shear measurement codes on simulated images, public challenges like the Shear TESting Programs (STEP) (Heymans et al. 2006; Massey et al. 2007) and the GRavitation lEnsing Accuracy Testing (GREAT) (Bridle et al. 2009; Kitching et al. 2012; Mandelbaum et al. 2014) were established. Valuable insight in the measurement process could be gained and significant progress was made. However, these challenges reaffirmed that a careful and rigorous treatment of systematic errors is essential to weak lensing as a cosmological probe.

Several large wide-field imaging surveys will come online in the next few years, including the Dark Energy Survey¹ (DES), the Kilo Degree Survey² (KiDS), Hyper Suprime-Cam³ (HSC), Euclid⁴, and the Large Synoptic Survey Telescope⁵ (LSST). They will map out a large fraction of the Sky, yielding a wealth of data. In this work, we will focus on data from DES.

In this paper, we present an initial implementation of a novel shear measurement approach, the *Monte Carlo Control Loops* (MCCL; Refregier & Amara 2014, henceforth RA13) at the one-point statistics level. We use image simulations that pass through the same lensing measurement pipeline as the data to forward model the measurement process. In this approach, not only can the shear measurement be calibrated, the nature of the

¹ <http://www.darkenergysurvey.org/>² <http://kids.strw.leidenuniv.nl/>³ <http://www.naoj.org/Projects/HSC/index.html>⁴ <http://sci.esa.int/euclid/>⁵ <http://www.lsst.org/lsst/>

† claudio.bruderer@phys.ethz.ch

pipeline allows us to test the robustness of the calibration. The *MCCL*-framework dynamically modifies the lensing pipeline and aims to provide a shear measurement with systematic errors smaller than the statistical errors for the survey being considered. For this purpose, we establish a set of three iterative *Control Loops (CL)* which build upon each other. First, the simulations are tuned to be statistically consistent with data. Second, the lensing measurement is calibrated. Third, the robustness of the calibration is tested. We achieve this by perturbing the simulation parameters and recalibrating the measurement while keeping data and simulations statistically consistent. The uncertainty in the calibration due to the perturbed parameters gives the systematic error of the measurement.

This paper is organized as follows. In Section 2, we explain the main concept and requirements in using the *MCCL* framework to tackle the shear measurement problem. In Section 3 we give a description of the DES SV data. The main features of UFIG are described in Section 4. We focus especially on the properties of the simulated galaxies, PSF, noise, and the shear field. In Section 5 we present the *MCCL* framework and its implementation. We show in Section 6 different diagnostics of our calibrated image simulations for DES. Furthermore, the results of our first tentative analysis of the robustness of the shear measurement calibration are presented. We conclude in Section 7.

2. *MCCL* AND THE SHEAR MEASUREMENT PROBLEM

The main goal of this paper is to tackle the weak lensing shear measurement problem using the *MCCL* approach proposed by RA13. In this section we elaborate on the main concepts behind the *MCCL* framework and how that translates into the specific implementations (see Section 5) carried out in this paper.

The first key concept is that all the *CLs* need to be specifically “controlled” by certain criteria, or targets. For example, in our first *CL*, we define criteria within which we view the simulations and the data to be statistically consistent. When the simulations satisfy the requirements, we leave the loop and continue on to the next step. The overall target that controls the entire framework is naturally tied to the science goals of the framework. In our case, the target is producing shear measurements that are accurate within statistical errors of the DES dataset of interest.

We choose to set the main target of this paper using results from AR08. First, we parametrize our measured shear to be linearly related to the true underlying shear via

$$\gamma_i^{obs} = (1 + m_i)\gamma_i^t + c_i + N, \quad (1)$$

where γ_i^t is the true shear, γ_i^{obs} the estimated shear. m_i and c_i are the multiplicative and additive biases, and N is a noise term that averages out for large numbers. Then, according to AR08, in order for the shear measurement not to be systematics-dominated, we would require $m_i \lesssim 0.025$ and $c_i \lesssim 1.65 \cdot 10^{-3}$ for a DES SV-like 200 deg² survey, and $m_i \lesssim 0.005$ and $c_i \lesssim 0.75 \cdot 10^{-3}$ for the full 5000 deg² DES survey. While these upper limits were derived by AR08 for two-point statistics, they place requirements on one-point statistics, namely that the absolute means of m_i and c_i must stay below the lim-

its stated. This can be thought of as the requirements for the case of spatially constant systematic effects.

Following the logic above, the second key concept is that the *MCCL* framework is problem- and survey-specific. The targets are set by the problem of interest, and the *CLs* are designed to achieve this sole target. This suggests that conclusions drawn from applying the *MCCL* approach should not be readily applied to different problems. For example, in this paper our goal is the measurement of shear one-point functions. Therefore, the results presented in this work are not appropriate to answer questions regarding two-point measurements of shear (e.g. spatial correlation of the shear measurements). A new *MCCL* framework with different target values and diagnostics will need to be designed for each particular question.

3. THE DARK ENERGY SURVEY

DES is a wide-field optical imaging survey that will cover 5000 deg² in the Southern Sky during its 5 years of operation and will record information of over 300 million galaxies. The survey area overlaps with other surveys such as the South Pole Telescope⁶ (SPT) and the Visible and Infrared Survey Telescope for Astronomy (VISTA) Hemisphere Survey⁷ (VHS). Focusing on Type Ia Supernovae, Baryon Acoustic Oscillations, galaxy clusters, and weak gravitational lensing as main cosmological probes, DES aims to study the nature of dark energy. The instrument achieved first light on 2012 September 12 and the main science survey officially started on 2013 August 12.

Images are taken with the Dark Energy Camera (DECam; Flaugher et al. 2012), designed specifically for DES. The camera is mounted on the Blanco-4m telescope at Cerro Tololo Inter-American Observatory in Chile. DECam provides 0.27"/pixel resolution. Good seeing on this site ranges between 0.7" and 1.1".

In this work, we will test our method using a fraction of the *SV-A1* release, which covers ~ 200 deg². Single exposures that were stacked to coadded images (straight averages), whose raw data are publicly available, were processed by the DES Data Management pipeline version “SVA1” (Yanny et al., in preparation). The images were taken during the SV period, which lasted between 2012 November and 2013 February. For this work, we selected images covering ~ 50 deg² in the SPT-E field that are free of significant image artifacts. We demonstrate our *MCCL* method on one image with an area of ~ 0.5 deg², *DES0441-4414*, while using the rest of this SPT-E subsample to derive the statistical errors. The area is sufficiently large and contains enough stars and galaxies for the simulations to be calibrated to this image.

4. ULTRA FAST IMAGE GENERATOR (UFIG)

In this paper we analyze images simulated with UFIG (Bergé et al. 2013, henceforth B13). The image generation process consists of two steps. First, galaxy and star catalogs are generated. Then, the catalogs are turned into a coadded image. A brief overview of the properties of the UFIG-generated galaxy catalogs, the PSF and noise models, and the shear field is given below, while

⁶ <http://pole.uchicago.edu/>

⁷ <http://www.vista-vhs.org/>

a full description can be found in B13. Note that some of the models used in B13 are not fully realistic, but they provide a good starting point. The output from our *MCCL* framework would inform us if more sophisticated models are needed to describe the data.

The *MCCL* approach typically requires the simulation and analysis of many thousands of images. Thus, speed is crucial. In order not to be dominated by the image generation, its speed needs to be at least comparable to the analysis. Due to several computational optimizations, UFig is orders of magnitude faster than publicly available image simulators. In terms of runtime, generating an image is comparable to executing *SEXTRACTOR* (Bertin & Arnouts 1996, henceforth *SE*) on the same image, which sets the time scale in the *MCCL* framework.

A key property of UFig is its flexibility in adjusting to different telescope setups. In this paper we choose to model *r*-band coadded images taken by DECam, but it is straightforward to simulate images from other wide-field imaging surveys.

4.1. Galaxies

A galaxy is simulated in UFig by sampling its light distribution photon-by-photon, and is then placed with a uniform probability on the image. Due to the finite number of photons sampled, the simulated galaxy images naturally include Poisson noise. PSF convolution in this approach is simply a displacement of the photons drawn from a probability distribution in the shape of the PSF (see Section 4.3). We model the galaxy light distribution with a single-Sérsic profile to which we apply a distortion to generate the apparent ellipticity.

The radial profile is defined by a Sérsic index, an intrinsic magnitude and an intrinsic size. The latter two are non-trivially correlated. We parametrize and sample this distribution in a space, where magnitudes and sizes are approximately uncorrelated (mag_r , $\log r_{50,r}$). It is related to the magnitude and size plane (mag , $\log r_{50,i}$) through a rotation by an angle θ around a pivotal point (mag_p , $\log r_{50,p}$), i.e.

$$\begin{pmatrix} mag \\ \log r_{50,i} \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} mag_r \\ \log r_{50,r} \end{pmatrix} + \begin{pmatrix} mag_p \\ \log r_{50,p} \end{pmatrix}. \quad (2)$$

We parametrize the distribution of rotated galaxy intrinsic half-light radii $r_{50,r}$ with a log-normal distribution with rms dispersion σ . The distribution of rotated magnitudes mag_r is approximated by the distribution of intrinsic magnitudes mag shifted by mag_p , which was compiled by B13 from different ground- and space-based surveys. In other words, we assume the cumulated magnitude distribution to be approximately invariant under the rotation described in Eq. 2. This is a good approximation for small values of the rotation angle θ . The two parameters θ and σ , the compiled cumulated magnitude distribution, and the pivotal point uniquely describe the two-dimensional distribution in the magnitude-size plane for our modeled galaxy sample. The Sérsic index distribution was derived by fitting single-Sérsic profiles to different galaxy samples in B13.

The intrinsic galaxy ellipticities are defined by (see e.g. Rhodes et al. 2000)

$$e = e_1 + ie_2 = \frac{I_{11} - I_{22} + 2iI_{12}}{I_{11} + I_{22}}, \quad (3)$$

where I_{ij} are the unweighted quadrupole moments of the galaxy's light profile and e_1 , e_2 are the two components of the ellipticity. In this paper, we sample e_1 and e_2 separately from normal distributions with mean zero and rms dispersion $e_{1,rms}$ and $e_{2,rms}$.

4.2. Stars

Since stars are typically brighter than galaxies, it is optimal to simulate them pixel-by-pixel rather than photon-by-photon. They are simulated directly on the image pixel grid and also placed on the image with a uniform probability. The profile is given by the PSF integrated within each pixel of the image grid (see Section 4.3). Poisson noise is included by drawing a value from the corresponding Poisson distribution in every pixel.

To simulate a star, only a magnitude needs to be drawn. We sample a cumulated magnitude distribution derived from the stellar population synthesis model Besançon (Robin et al. 2003). In case the resulting intensity in a pixel is larger than DECam's saturation threshold, bleeding is modeled.

4.3. PSF

In this initial implementation of the *MCCL* framework we choose as a baseline for future work a spatially constant, elliptical Moffat profile to describe the PSF. The elliptical Moffat profile can be derived from a two-dimensional linear transformation of the circular Moffat profile, which is given by (Moffat 1969)

$$I(r) = \frac{I_0}{\left(1 + \left(\frac{r}{\alpha}\right)^2\right)^\beta}, \quad (4)$$

where the scale parameter α is related to the seeing. The profile is defined by the seeing, the exponent β , and the ellipticities e_1 and e_2 . We find that the radial profile of stars in coadded DES images roughly follow a Moffat distribution with some variation in the parameters α and β .

In this initial implementation we choose for simplicity to fit a spatially invariant PSF to the image of interest (*DES0441-4414*, see Section 3) in a pre-calibration step. We use in this work a PSF of FWHM 1.09'', ellipticity $e_1 = 0.035$, $e_2 = 0.02$, and $\beta = 3.5$ to match the mean PSF of this image. Note that this PSF size is slightly larger than the projected median seeing of the main survey ($\sim 0.9''$). As shear measurement is more challenging with larger PSF sizes, we expect our *MCCL* framework to produce results similar or better on an image with better seeing conditions.

4.4. Noise

Two different components add to the noise in UFig. First, we simulate galaxies down to magnitudes $r \sim 29$. Since most of these faint galaxies are not detected, every image contains sky noise arising from many unresolved, faint galaxies. Second, we add a Gaussian background noise centered around 0 with a constant rms dispersion σ_N across the image. This should capture noise induced by emission from the sky, and noise induced by the data processing. We perform furthermore Lanczos resampling (Duchon 1979) with a kernel of width five pixels and a

half-a-pixel offset on the simulated pixel grid. It allows us to mimic correlated noise in real images, while bypassing the expensive simulation and data reduction of raw images (see B13).

4.5. Shear field

We employ the following shear conventions in UFIG and throughout this paper (Rhodes et al. 2001; Bartelmann & Schneider 2001)

$$\gamma_1 = \frac{1}{2} (\partial_1^2 - \partial_2^2) \Psi \quad \text{and} \quad \gamma_2 = \partial_1 \partial_2 \Psi, \quad (5)$$

where Ψ is the projected lensing potential.

To be close to real surveys, we use a Λ CDM shear power spectrum and model the shear field as a Gaussian random field. We choose $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\sigma_8 = 0.8$. We simulate Gaussian random fields with Lang & Potthoff (2011)'s fast algorithm.

5. METHOD

The *MCCL* framework is designed to validate the shear measurement process on simulated images and to test its robustness. RA13 identify three key iterative steps in the shear measurement process, which are labeled as *Control Loops* (*CL*), each with a distinct goal.

The first step (*CL1*) is designed to find a fiducial configuration of simulation parameters such that the simulations agree with the data. In order to quantify the level of agreement, this step relies on defining a set of diagnostics and metric targets. The next step (*CL2*) is to calibrate the shear measurement at this fiducial point. The final and computationally most demanding step (*CL3*) aims to explore the parameter space volume for which data and simulations are in good agreement to ensure that the calibration scheme from *CL2* is robust. This scheme is designed to ensure that the systematic errors on a given shear measurement are sub-dominant to the statistical errors. Should the results of *CL3* show that the employed calibration scheme is not robust enough over all parameter space allowed by *CL1*, then the whole *MCCL* framework needs to be applied again with more stringent diagnostic requirements and possibly additional diagnostics.

It is clear now that in order to assess the robustness of this calibration scheme the generation and analysis of many tens or even hundreds of thousands of images is required. From a computational viewpoint this is only feasible if every step is very fast. This echoes our statement in Section 4 on the importance of using UFIG as our main image simulation tool.

The detailed implementation of each of the *CLs* is presented below.

5.1. Control Loop 1

To make statements about the consistency of data and simulations output, we analyze three distributions described below as our main diagnostics. To assess how likely it is that two different distributions of data and simulations could be different realizations of the same underlying model, we use a χ^2 -method. We apply appropriate cuts to the three diagnostic distributions and bin them. This allows us to compute χ^2 for each diagnostic and combine them by adding them up. For a number

of different diagnostic distributions $\# \text{Diag}$, the total χ_{red}^2 for two binned datasets of different sizes is given by (e.g. Press et al. 2002)

$$\chi_{red}^2 = \frac{1}{\sum_k N_k} \sum_{i=1}^{\# \text{Diag}} \sum_{j=1}^{N_i} \frac{\left(\sqrt{g_i/f_i} f_{ij} - \sqrt{f_i/g_i} g_{ij} \right)^2}{\sigma_{d,ij}^2 + \sigma_{s,ij}^2}. \quad (6)$$

Here, for the i -th diagnostic distribution of the real (simulated) image, f_{ij} (g_{ij}) is the number of counts in the j -th bin. N_i bins is the number of bins for this diagnostic with counts f_{ij} above a certain threshold, and f_i (g_i) is the sum of all counts in those bins. $\sigma_{d,ij}$ and $\sigma_{s,ij}$ are the errors of the data and the simulation for the i -th diagnostic distribution within the j -th bin. The errors need to be estimated in the data and the simulations. For the data, we estimate them by computing the variance in those bins for all the images in the sample. For the simulations, we generate many different realizations of the same input model and compute the variances in every bin. For this χ^2 -method the variables in each bin should follow a Gaussian distribution. We therefore only include bins with at least 50 objects (about 36000 objects are detected in the real image). We find this to be a good approximation.

For this first implementation, we choose three diagnostics to break the degeneracies between the parameters we vary. They are refined iteratively to meet the requirements to pass *CL3* (see Section 5.3). Two of the three diagnostics use SE estimators. A comparison of the performance of certain estimators on these images is shown in (e.g. Bertin & Arnouts 1996; Chang et al. in preparation, 2014, and references therein). The three diagnostics are:

- Histogram of pixel values in ADUs (Fig. 2): 1D

This is a valuable diagnostic to test the background properties of the image by comparing the peak of the distribution in the sky-subtracted images. Furthermore, it allows us to test the magnitude zero point of the image. Different magnitude zero points shift the tail of the large pixel values vertically, as they affect the number of pixels with small respectively large pixel values.

- Binned magnitude versus size-plane (Fig. 3): 2D

This diagnostic probes the magnitude and size distributions of identified objects in the images and their correlation. We use the SE columns MAG_BEST for the magnitude and FLUX_RADIUS for the size.

- Binned e_1 versus e_2 -plane in three different magnitude bins (Fig. 4): 2D

This tests the ellipticity distribution of identified objects. We estimate the ellipticity using a version of Eq. 3 with weighted quadrupole moments. We split the objects up into three different magnitude bins, each containing a similar number of objects. This allows us to probe the ellipticity distribution in each S/N bin individually. With the high-S/N bin being least affected by the effects of the PSF, different intrinsic ellipticity distributions

can be distinguished. The low-S/N bin, which contains faint objects whose shape is dominated by the PSF, on the other hand allows us to test the properties of the PSF.

We minimize χ_{red}^2 to find a fiducial configuration. For this first implementation, we choose to vary six simulation parameters to generate new samples describing the galaxy population, the image properties, and the noise level: the magnitude zero point of the image mag_0 , the rms of the log-normal size distribution σ (Section 4.1), the rotation angle θ between the magnitude and size plane and the plane where the quantities are approximately uncorrelated (Section 4.1), the rms of the Gaussian background noise σ_N (Section 4.4), and the rms of the Gaussian distributions for the ellipticities $e_{1,rms}$ and $e_{2,rms}$ (Eq. 3). Those six parameters are not constrained by fits performed in B13. For each configuration an image is simulated and the χ_{red}^2 -value is computed (Eq. 6).

The χ_{red}^2 minimization procedure is designed to find a sensible parameter regime in a small number of iterations. It consists of two steps: First, we sample the parameter space coarsely and identify the region in the parameter space where the minimum χ_{red}^2 is located. Then, by successive one-dimensional minimizations we find the minimum χ_{red}^2 . We vary each parameter while holding the others fixed and compute the new χ_{red}^2 values. The specific parameter value that minimizes χ_{red}^2 defines a new configuration. We repeat this step iteratively until it converges. The final result of the iteration is the fiducial configuration that is analyzed in the subsequent CLs.

If Eq. 6 can be applied, i.e. the quantities in each bin of the diagnostic distributions are Gaussian distributed, then confidence limits on the parameters can be computed. In this case, for a model with six degrees of freedom the 95% confidence limits are given by (e.g. Chernick & Friis 2003)

$$\Delta\chi_{red}^2 = \frac{1}{\sum_k N_k} \cdot 12.59. \quad (7)$$

This gives for every parameter a range of values for which data and simulations are statistically consistent (see Appendix).

5.2. Control Loop 2

The task of *CL2* is to calibrate the shear measurement by comparing input and estimated shear signal on simulated data. The image that we use to illustrate the *MCCL* framework is part of the DES *SV-A1* release, which covers about 200 deg². The larger the area simulated for calibration is, which needs to be larger than the size of the dataset, the more precise is the resulting calibration scheme. On the other hand however the computational costs increase for a larger area. We choose to simulate an area equivalent to 1000 deg² for any given configuration. On the galaxies detected in those images we apply a S/N-cut of 15, where we define the S/N as SE's FLUX_BEST/FLUXERR_BEST, and a size-cut of 1.2 times the PSF size using SE's FLUX_RADIUS measurement. This allows us to select galaxies large and bright enough for calibration.

We follow Rhodes et al. (2001) to first order to estimate the galaxy shear

$$\hat{\gamma} = \frac{e'}{2 - \langle |e|^2 \rangle}, \quad (8)$$

where

$$e' = \frac{J'_{11} - J'_{22} + 2iJ'_{12}}{J'_{11} + J'_{22}} \quad (9)$$

is the lensed ellipticity, and e the unlensed one. In the weak lensing limit we can approximate

$$\langle |e|^2 \rangle \approx \langle |e'|^2 \rangle. \quad (10)$$

We use SE's X2WIN_IMAGE, Y2WIN_IMAGE, and XYWIN_IMAGE to measure the weighted quadrupole moments of the PSF-convolved image \tilde{J}'_{ij} . To linear order and ignoring weight function terms, the PSF can approximately be corrected for using

$$J'_{ij} = \tilde{J}'_{ij} - P_{ij}, \quad (11)$$

where P_{ij} is the mean of the weighted quadrupole moments of the stars. We use $J'_{ij} = \tilde{J}'_{ij}$ in Eq. 9 when PSF correction is not applied.

The galaxies are then binned in input shear signal and the mean estimated shear is computed in every bin. We calibrate the shear measurement to first order by fitting and applying a linear correction

$$\hat{\gamma}_i = \alpha_i \gamma_{in,i} + \beta_i, \quad (12)$$

where $\gamma_{in,i}$ is the input shear.

5.3. Control Loop 3

Knowing the ranges of parameter values for which data and simulations are statistically consistent from *CL1*, we can test the robustness of the calibration schemes for shear measurements for different configurations in this parameter space volume (*CL3.1*). We vary each parameter in a range slightly larger than that allowed by the data, while keeping the other parameters fixed, and calibrate the shear measurement on this new location in parameter space (see Section 5.2). We then explore by which amount the calibration (α and β ; see Eq. 12) changes relative to the calibration on the fiducial configuration resulting from applying *CL1*. This uncertainty in the shear calibration corresponds to the systematic error we expect in the shear measurement.

The resulting multiplicative and additive biases (Eq. 1) due to the uncertainty in the fiducial configuration are computed by evaluating

$$m_i = \frac{\Delta\alpha_i}{\alpha_i} \quad \text{and} \quad c_i = \frac{\Delta\beta_i}{\alpha_i}, \quad (13)$$

where $\Delta\alpha_i$ and $\Delta\beta_i$ are the changes relative to the fiducial calibration parameters. We require m_i and c_i to meet the targets set in Section 2, otherwise the diagnostics themselves need to be refined and additional tests could be required (*CL3.2*), affecting all the previous loops.

6. RESULTS

6.1. Control Loop 1



Figure 1. Comparison of a DES SV image (*DES0441-4414*; left) and a UFIG simulated image after *CL1* (right). A 4 arcmin² segment of the total 0.5 deg² images is shown. The same color scale has been applied to both images.

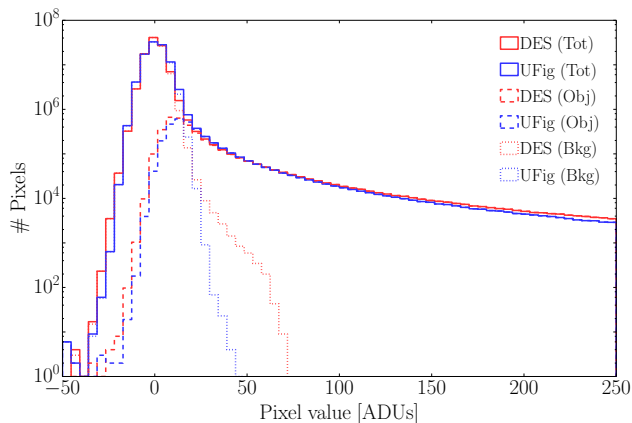


Figure 2. Histogram of the pixel values in ADUs for *DES0441-4414* (red) and a simulated UFIG image (blue) with the fiducial configuration after *CL1*. The solid lines show the counts of all the pixels in the image. SE’s segmentation map assigns pixels either to objects or the background. The dashed respectively dotted lines show the corresponding pixel counts.

Excerpts of the *DES0441-4414* image and the UFIG image simulated with the fiducial configuration are displayed in Fig. 1. They appear similar visually. For a quantitative comparison, Figures 2-4 show the diagnostic plots for the DES image and the UFIG image. The combined χ^2_{red} of the individual values for each diagnostic has a value of 1.06. Thus, the fiducial configuration we find is a good fit to the data in the chosen diagnostics. To avoid combining very different χ^2_{red} values, we assure that the individual χ^2_{red} values are also close to 1. For the fiducial configuration, the individual ones for each diagnostic are within $|\chi^2_{red} - 1| < 0.4$ (see Appendix). By varying the binning scheme we have checked that we recover similar fiducial configurations and confidence limits.

Fig. 2 shows the histograms of pixel values for all the

pixels in both images (solid). The overall behavior agrees well ($\chi^2_{red} \approx 1.38$). The histograms agree well around the peak, with the distribution of the pixels in the UFIG image being slightly broader. The pixels are furthermore divided using SE’s segmentation map into two sets to allow us to understand differences and similarities better. One set contains all the pixels associated with identified objects (dashed), and the other those associated with the background (dotted). The histograms of pixels associated with objects agree well ($\chi^2_{red} \approx 1.10$). We however observe a low-level discrepancy in the background pixel histograms at high pixel values. While our noise model including Gaussian noise in every pixels seems to be a good approximation around the peak of the histogram, it does not account for the background pixels with larger positive pixel values. As the number of background pixels is small compared to the total number of pixels with pixel values of $\gtrsim 30$ ADUs, those differences do not affect the value of χ^2_{red} significantly.

Fig. 3 displays the magnitude-size plane of objects identified by SE in both the simulation and the data. Overall, the distributions resemble each other qualitatively and quantitatively ($\chi^2_{red} \approx 1.26$). In particular, the main bulk of the galaxy distributions, the location of the stellar loci, and the saturation turnoffs all agree well. Some slight differences can however be noted. The dispersion around the stellar locus is larger in the DES image, which is due to our simple PSF model constant in size. Furthermore, the shapes of the density contour lines and the magnitude limits are slightly different. We believe that changes in the galaxy model would improve this.

The different magnitude limits and the discrepancies in the background-only histograms of pixel values call for more noise in the simulations. Increasing the width of the Gaussian background would on the other hand how-

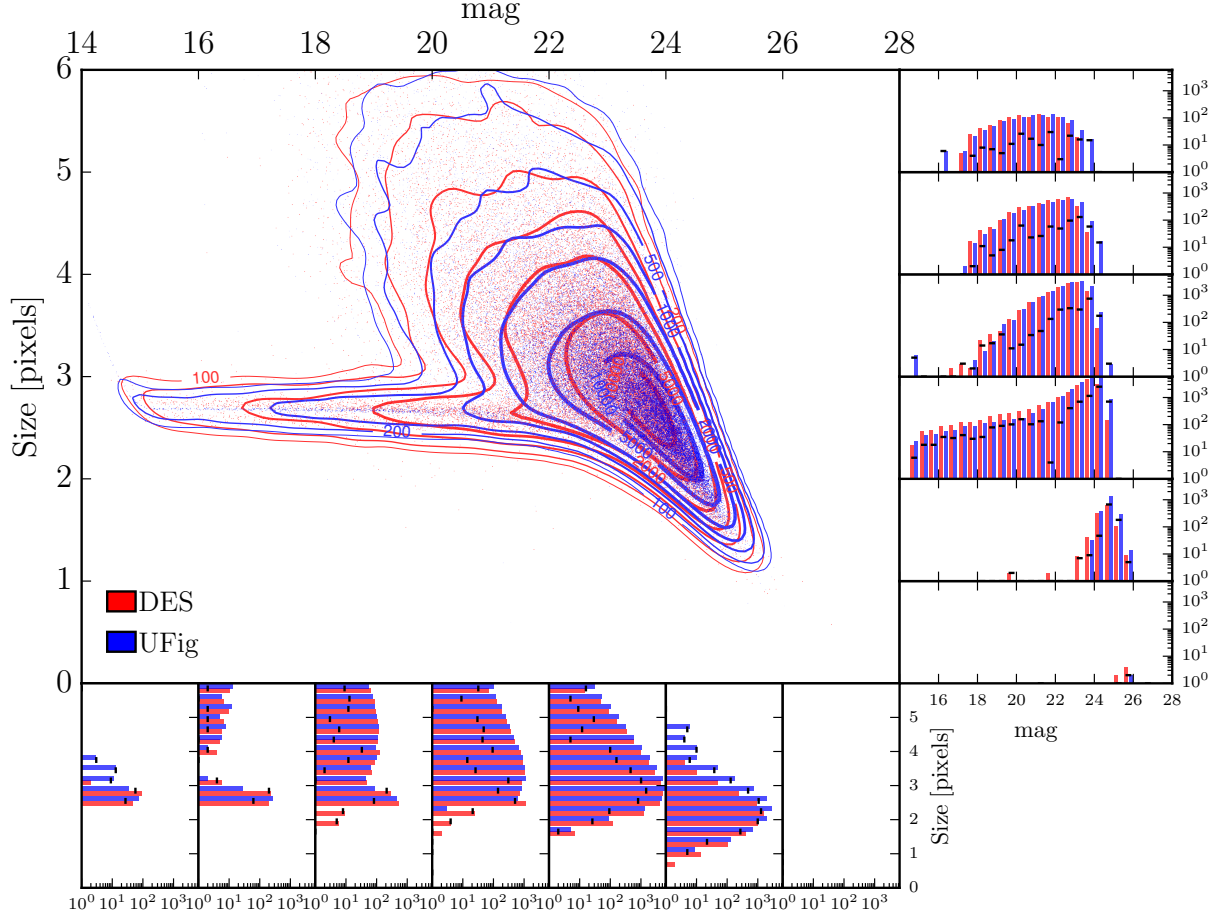


Figure 3. Distribution of r -band magnitudes (MAG_BEST) and the sizes in pixels (FLUX_RADIUS) of objects identified by SE. Isodensity contours of the number of objects track the shape of the distribution. Red is the *DES0441-4414* and blue is a simulated UFig image with the fiducial configuration after *CL1*. Histograms on the right and the bottom show the projected distributions in different size and magnitude bins. The black marks denote the difference between the red and blue histograms in every bin.

ever aggravate the discrepancy around the background peak. To resolve this tension (see Appendix), a more sophisticated background model easing some of the simplifying assumptions on the properties of the background is needed (for an overview of possible extensions see Rowe et al. 2014). An analysis of the two-point correlation function will reveal structures in the background not yet modeled and will serve as an additional diagnostic.

The ellipticity planes in the different magnitude bins are shown in Fig. 4. Due to the ellipticity introduced by the PSF, the mean of the e_i -distributions is shifted towards positive values and thus there is a small asymmetry. Note that the galaxies we include in the calibration of the shear measurement are mainly in the two brighter magnitude bins where the distributions match well ($\chi^2_{red} \approx 1.35$ and $\chi^2_{red} \approx 0.63$). In the brightest magnitude bin, the distributions deviate slightly for values of $|e_i| > 0.3$. We believe this is caused by our choice of the intrinsic ellipticity distributions being normal in e_i (see Eq. 3). Changes in the intrinsic ellipticity distribution can improve the agreement between the data and the simulation. In the faintest magnitude bin, the distributions do not match well ($\chi^2_{red} \approx 0.21$). As noted

above, there seem to be more faint objects detected in the UFig image (about $\sim 15\%$ more detected objects in total). Furthermore, there are differences between the ellipticity distributions in this bin. This can be attributed to the simple PSF model we choose, as the objects in this bin are mostly dominated by the PSF. As we apply a S/N-cut of 15 (corresponds to $\text{mag} \sim 23$), differences in the faintest magnitude are potentially not relevant for the calibration of the shear measurement. Nevertheless, it is only by looking at the results of a future, more rigorous *MCCL* analysis including parameters describing the PSF model that we can assess whether the differences in the faintest magnitude bin are relevant for shear measurement.

6.2. Control Loops 2 and 3

We perform a tolerance analysis of the shear calibration, as described in Section 5.3. We vary the same six parameters as in Section 5.1, mag_0 , σ , θ , σ_N , $e_{1,rms}$, and $e_{2,rms}$. The allowed parameter ranges by the data are given by the analysis performed in Section 6.1 (see Appendix). They correspond to the 95% confidence limits we take as a measure of statistical consistency of dif-

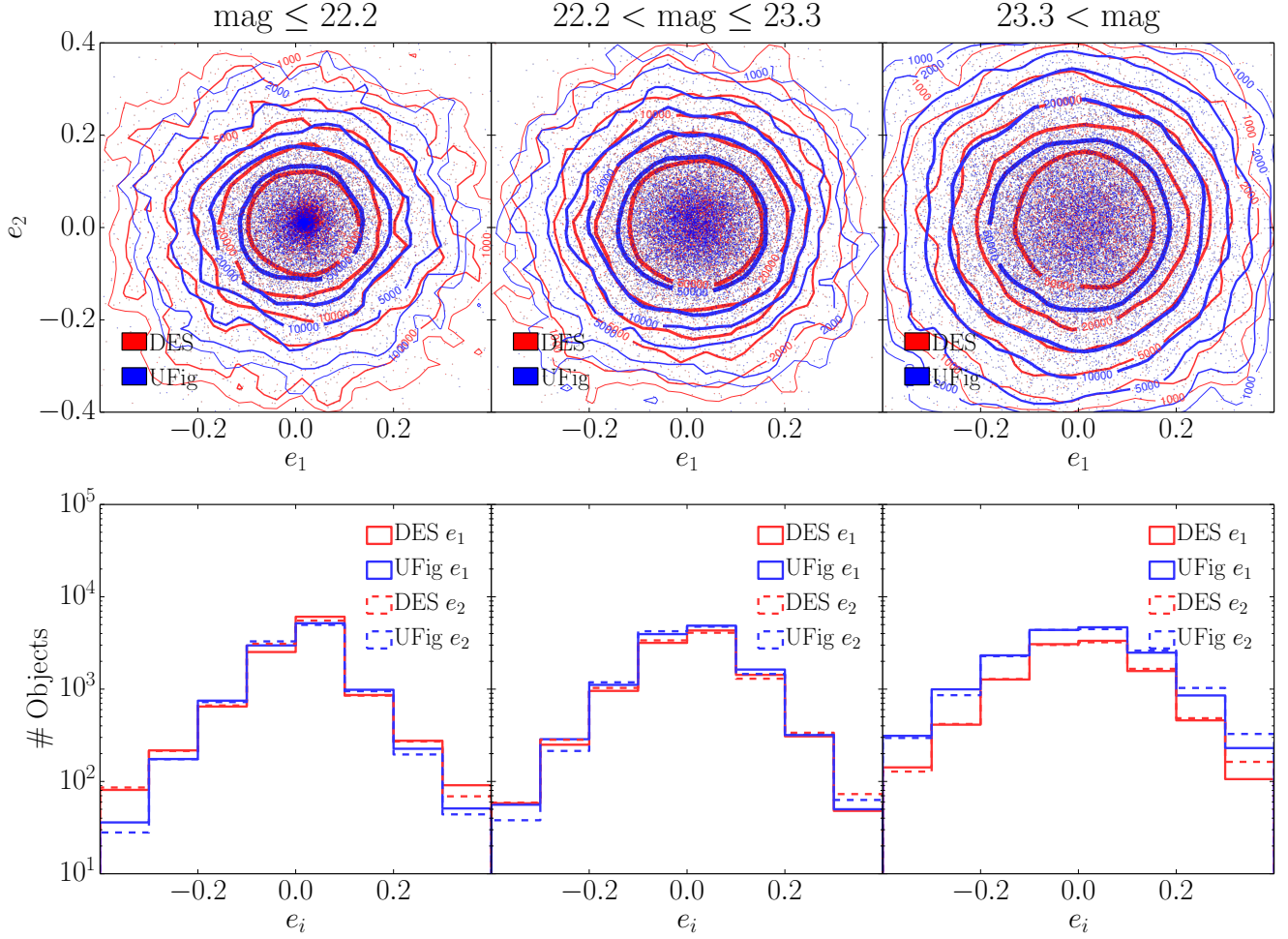


Figure 4. Ellipticity distributions for objects identified by SE in *DES0441-4414* (red) and a simulated UFig image with the fiducial configuration after *CL1* (blue) are shown. The objects are split up into three different *r*-band magnitude (*MAG_BEST*) bins, such that they contain approximately the same number of objects. In every bin, isodensity contours of identified objects in the ellipticity plane (see Eq. 3) (top row) and the corresponding histograms (bottom row) are shown.

Table 1
Calibration coefficients α_i and β_i for the fiducial configuration

Coefficient	PSF-uncorrected	PSF-corrected
α_1	0.199 ± 0.001	0.380 ± 0.002
α_2	0.197 ± 0.001	0.377 ± 0.002
β_1	$(4.40 \pm 0.01) \cdot 10^{-3}$	$(0.82 \pm 0.02) \cdot 10^{-3}$
β_2	$(2.86 \pm 0.01) \cdot 10^{-3}$	$(1.10 \pm 0.03) \cdot 10^{-3}$

ferent configurations. For each of these parameters, we compute the change in calibration relative to the fiducial model (see Table 1) at six different points around the fiducial configuration. For every data point we simulate an area of 1000 deg^2 . Thus, to calibrate the shear measurement with the precision required for 200 deg^2 survey, we need to simulate 37000 deg^2 .

Figures 5 and 6 show how uncertainties in the input parameters result in multiplicative biases. We use the two shape measures described by Equations 8-11. We find that, for the six parameters we vary, the PSF-corrected shape measurement calibrated through the *MCCL* framework is robust enough for a DES SV-like 200 deg^2 sur-

vey in terms of the requirement described in Section 2. As discussed in (Refregier & Amara 2014) unknown systematics or effects not yet included in the simulations may affect the shear measurement. However, the *MCCL* approach provides a framework for testing aspects of the measurement process that are in doubt. The PSF-uncorrected shape measurement does not perform as well as the PSF-corrected one, and lies slightly outside the tolerance band in some parameters. To make statements about whether the calibration scheme is robust enough for a 5-year DES-like 5000 deg^2 -survey in the parameters varied, a larger area needs to be simulated to increase the accuracy of the calibration. Furthermore, as described in Section 2, achieving this new target requires refinements on the *MCCL* framework.

Figures 7 and 8 show the resulting additive biases. For the parameters considered, both shape measures already even satisfy the requirements for a full DES-like survey with 5 years worth of images.

We find in this first tolerance analysis that the calibration of the shear measurement seems to depend sensitively on the intrinsic ellipticity distribution. While there is not a significant additive bias due to an uncertainty in

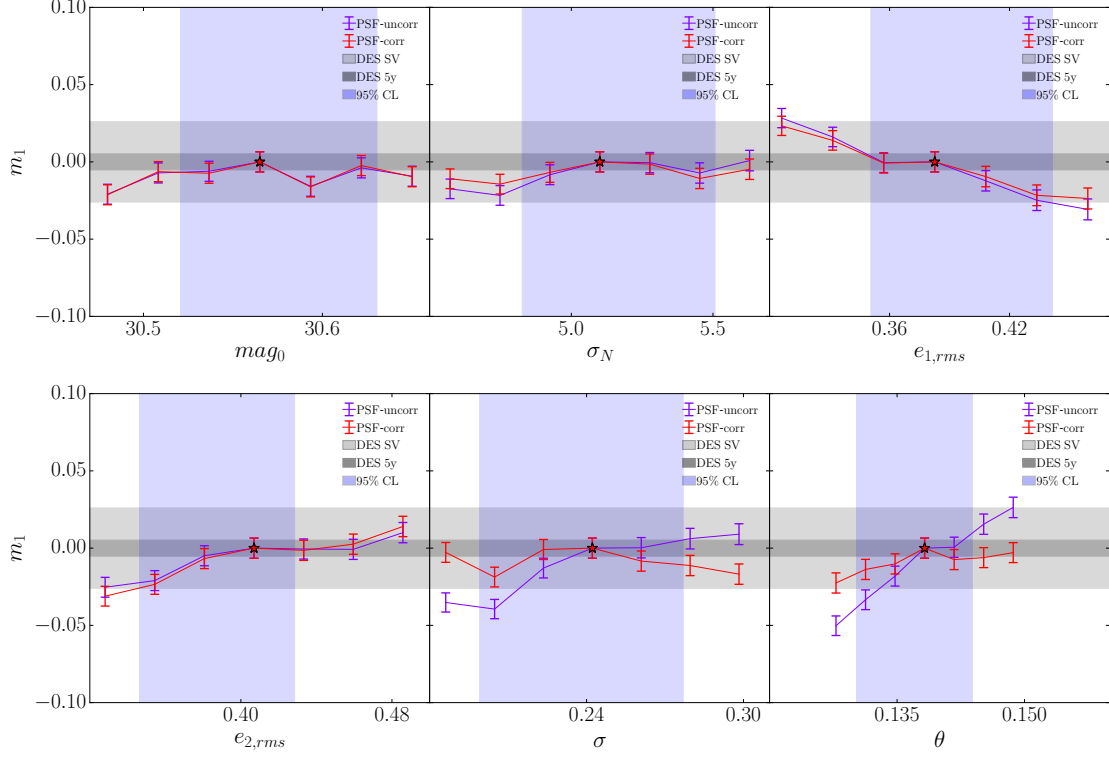


Figure 5. Multiplicative bias in the measurement of γ_1 as a function of different parameter values and different shear measurement methods. We simulate images equivalent to an area of 1000 deg^2 for every configuration to calibrate the shear measurement. The change is relative to the central data point, our fiducial shear calibration. The vertical blue bands show the range in parameter values data and simulations are statistically consistent (95% confidence limits) (see Section 4.1). The horizontal gray bands correspond to the required accuracy in the shear measurement of a 200 deg^2 (light gray) and 5000 deg^2 survey for the measurement not to be systematics-limited. The star denotes the fiducial configuration.

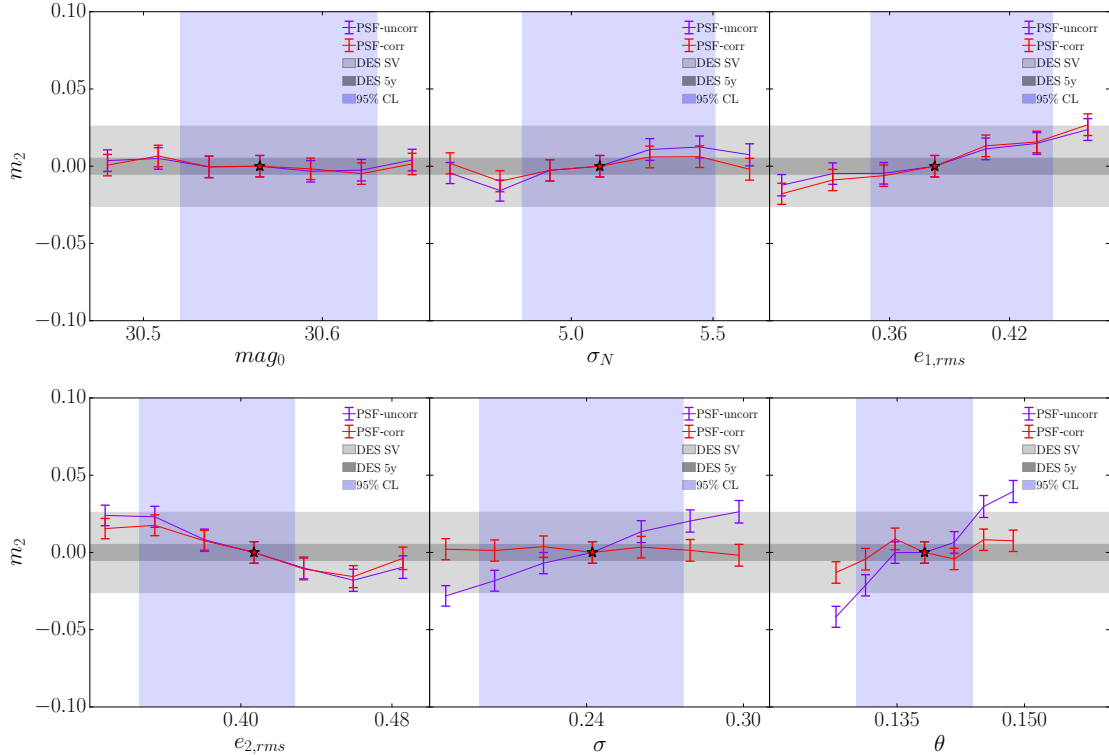


Figure 6. Multiplicative bias in the measurement of γ_2 . Similar to Fig. 5.

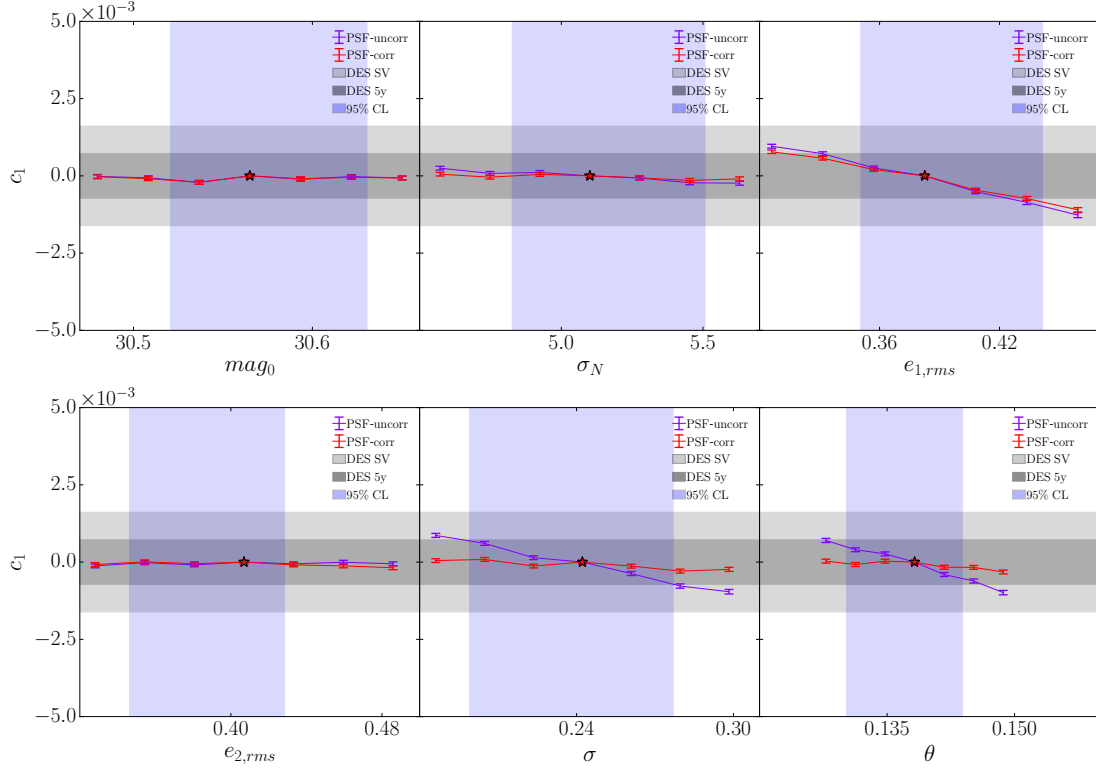


Figure 7. Additive bias in the measurement of γ_1 . Similar to Fig. 5.

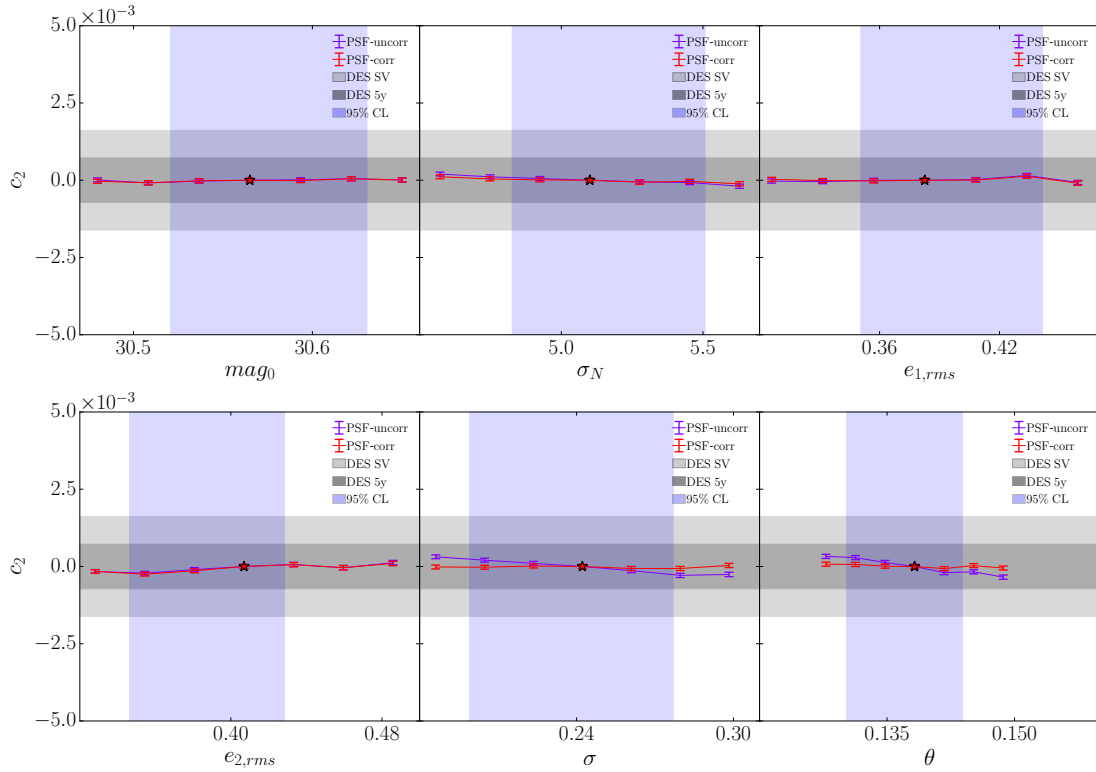


Figure 8. Additive bias in the measurement of γ_2 . Similar to Fig. 7.

$e_{1,rms}$ and $e_{2,rms}$, the ellipticity distribution needs to be taken special care of such that no significant multiplicative bias is induced. The diagnostics likely need to be refined further to reduce this residual systematic effect such that stricter targets can be met in further *MCCL* analyses.

7. CONCLUSION

We have presented an initial implementation of the *Monte Carlo Control Loops* (Refregier & Amara 2014), a novel approach for weak lensing shear measurements. The method contains a set of three *Control Loops* (*CL*) applied to data and image simulations to forward-model the shear measurement process. They are designed specifically to calibrate the shear measurement and test its robustness with the goal of reaching a certain sensitivity. The requirements in this paper are chosen such that the lensing measurement, assuming spatially invariant systematic errors, on the final dataset of a DES- and also DES SV-like imaging survey is not limited by systematic errors, i.e. the systematic error of the measurement is smaller than the statistical error.

The *MCCL* approach provides a consistent way of analyzing systematic errors in the measurement. It allows us to probe potential sources of error for their effect on the measurement, e.g. noise bias (Refregier et al. 2012; Kacprzak et al. 2012) and model bias (e.g. Kacprzak et al. 2014), provided that they are included in the simulations. However, the simulation and analysis of a large number of images is essential in this approach. To not be limited computationally, every step in the *CLs* needs to be fast, especially the generation of images. This led to the development of the Ultra Fast Image Generator (Bergé et al. 2013), whose speed is comparable to executing SEXTRACTOR (Bertin & Arnouts 1996), the image analysis tool used in this paper.

We present a first implementation of the *MCCL* framework using an image taken during the Science Verification (SV) phase of DES. For this purpose, we choose a spatially invariant PSF model, vary six simulation parameters, and consider only one-point shear measurements. With these assumptions, we find that the image calibration achieves multiplicative and additive biases within the needed weak lensing precision for a DES SV-like (200 deg²) survey, assuming them to be spatially invariant. We also find with the tolerance analysis that the shear measurement is very sensitive to the intrinsic ellipticity distribution. Furthermore, we find an interplay between the magnitude-size and the histogram of pixel values diagnostics in fitting the noise level to the image. To accommodate both diagnostics, an extension of the Gaussian noise model will be implemented in future work.

To achieve our goal of not being systematics-limited when measuring shear on a 5000 deg² DES-like survey, several features in the *MCCL* framework need to be refined. First, we will incorporate more realistic instrument and noise model in the image simulations. Next, we like to extend this framework to include two-point functions in the analysis. In addition, we are planning to test the effects of a spatially varying PSF and other PSF models. We will also explore the effect of more complex galaxy models and non-uniform distributions of galaxies on the calibration of the shear measurement. And

finally, a more rigorous tolerance analysis varying more simulation parameters is required.

The results we present in this work require the simulation of about 40000 deg² of images. From a simple extrapolation of this figure, the computational resources needed for the full 5-year DES data appear large. However, several improvements on the framework can readily result in significant speed-ups. For example, better sampling strategies can in principle speed up the tolerance analysis in *CL3.1* by orders of magnitude, which is by far the most computationally expensive step in this work. Furthermore, improvements in the diagnostics used in *CL1* will increase the discriminatory power between different simulation configurations, reducing further the parameter space one needs to sample, and thus the computational time.

All these improvements will pave the way to exploit the full potential of weak lensing through the understanding of systematic effects within the *MCCL* framework.

8. ACKNOWLEDGEMENTS

The authors would like to thank Sarah Bridle, Tomasz Kacprzak, David Bacon, Matthew Becker, Barnaby Rowe, Gary Bernstein, and the other members of the DES collaborations for useful discussion. This work was supported in part by grants 200021.14944 and 200021.143906 from the Swiss National Science Foundation.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência e Tecnologia, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Edinburgh, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

REFERENCES

- Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, ArXiv Astrophysics e-prints, astro-ph/0609591
- Amara, A., & Réfrégier, A. 2008, MNRAS, 391, 228
- Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291
- Bergé, J., Gamper, L., Réfrégier, A., & Amara, A. 2013, Astronomy and Computing, Volume 1, p. 23-32., 1, 23
- Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
- Bridle, S., Shawe-Taylor, J., Amara, A., et al. 2009, Annals of Applied Statistics, 3, 6
- Chang, C., Busha, M. T., Wechsler, R. H., et al. in preparation, 2014
- Chernick, M., & Friis, R. 2003, Introductory Biostatistics for the Health Sciences (Wiley, New York)
- Duchon, C. E. 1979, Journal of Applied Meteorology, 18, 1016
- Flaugher, B. L., Abbott, T. M. C., Angstadt, R., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
- Hoekstra, H., & Jain, B. 2008, Annual Review of Nuclear and Particle Science, 58, 99
- Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, MNRAS, 366, 101
- Kacprzak, T., Bridle, S., Rowe, B., et al. 2014, MNRAS, 441, 2528
- Kacprzak, T., Zuntz, J., Rowe, B., et al. 2012, MNRAS, 427, 2711
- Kitching, T. D., Balan, S. T., Bridle, S., et al. 2012, MNRAS, 423, 3163
- Lang, A., & Potthoff, J. 2011, ArXiv e-prints, arXiv:1105.2737
- Mandelbaum, R., Rowe, B., Bosch, J., et al. 2014, ApJS, 212, 5
- Massey, R., Heymans, C., Bergé, J., et al. 2007, MNRAS, 376, 13
- Moffat, A. F. J. 1969, A&A, 3, 455
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical recipes in C++ : the art of scientific computing
- Refregier, A. 2003, ARA&A, 41, 645
- Refregier, A., & Amara, A. 2014, Physics of the Dark Universe, 3, 1
- Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. 2012, MNRAS, 425, 1951
- Rhodes, J., Refregier, A., & Groth, E. J. 2000, ApJ, 536, 79
- . 2001, ApJ, 552, L85
- Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523
- Rowe, B., Jarvis, M., Mandelbaum, R., et al. 2014, ArXiv e-prints, arXiv:1407.7676
- Zuntz, J., Kacprzak, T., Voigt, L., et al. 2013, MNRAS, 434, 1604

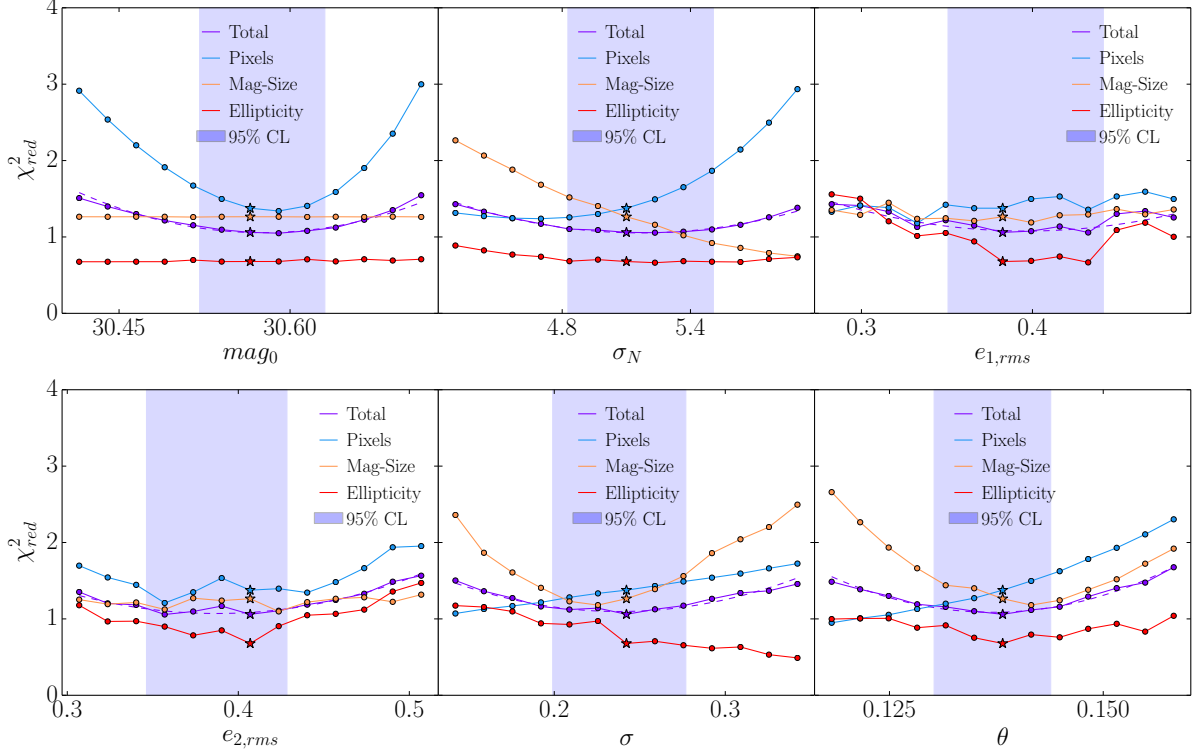


Figure 9. χ^2_{red} -curves for the three different diagnostics as a function of different parameter values are shown. The star denotes the fiducial configuration. The vertical blue bands are computed with Eq. 7 and show the 95%-confidence limits. They correspond to the blue bands in Figures 5-8.

APPENDIX

REDUCED χ^2 AS A FUNCTION OF DIFFERENT SIMULATION PARAMETERS

As described in Section 5.1, we search for the fiducial simulation configuration by minimizing the χ^2_{red} defined in Eq. 6. In this Appendix, we describe in detail the minimization procedure and point out some features in the resulting χ^2_{red} functions, which may indicate interesting physical insights to the data.

For each of the six simulation parameters considered, we systematically vary its value around some initial guess and calculate χ^2_{red} while holding the other parameters fixed. The six parameter values that yield the minimum χ^2_{red} are then used for the next iteration and the process continues until it converges about the minimum. In this final set of parameters, the χ^2_{red} values along each of the one-dimensional axes are shown in Fig. 9. The blue shaded bands, which correspond to the blue bands in Figures 5-8, are the 95%-confidence limits for each of the parameters (see Eq. 7). Table 2 lists the corresponding parameter values in the plots.

To better understand how the various diagnostics affect the resulting χ^2_{red} -function, we split it up into the contributions of each diagnostic. For the fiducial configuration, which is denoted by a star, we find $|\chi^2_{red} - 1| < 0.4$ for all the individual diagnostics and their combined sum. The total χ^2_{red} is well approximated by quadratic fits, though χ^2_{red} from individual diagnostics can show very different behaviors. Furthermore, the fiducial configuration is close to the minima of the quadratic fits.

We want to point out a few features in the individual subfigures. First, the magnitude-size plane (see Fig. 3) and the ellipticity plane (see Fig. 4) do not react to changes in the magnitude zero point of the image (mag_0). Only the histogram of pixel values (see Fig. 2) responds to changes in mag_0 , as the magnitude zero point affects the normalization of the pixel values.

Second, when varying the width of the Gaussian background σ_N , the histogram of pixel values and the magnitude-size plane respond in opposite directions. As described in Section 6.1, the fiducial model produces a slightly deeper image compared to data. Thus, the magnitude-size plane is pushing for a higher noise level. The histogram of pixel values however constrains the background peak in the sky-subtracted image and cannot accommodate larger σ_N values. Therefore, we believe that including an additional Non-Gaussian noise term can reconcile the tension.

Third, the diagnostics are rather flat if $e_{1,rms}$ and $e_{2,rms}$ are varied. As a result, the 95%-confidence limits for this parameter are relatively wide. Another less noisy diagnostic to constrain the ellipticity distribution might shrink these confidence limits.

Table 2
Minimum and 95%-confidence limits of the quadratic fits (Fig. 9) and parameter values for the fiducial configuration

Parameter	Fiducial value	Central value	95% CL	Description
mag_0	30.565	30.576	± 0.055	Magnitude zero point
σ_N	5.10	5.17	± 0.34	rms of the background noise
$e_{1,rms}$	0.383	0.396	± 0.045	rms of the intrinsic e_1 distribution
$e_{2,rms}$	0.407	0.387	± 0.041	rms of the intrinsic e_2 distribution
σ	0.2422	0.2381	± 0.0390	rms of the intrinsic log-normal size distribution
θ	0.1382	0.1370	± 0.0068	Rotation angle to plane intrinsic magnitudes and sizes are uncorrelated

Fourth, the χ_{red}^2 values of the ellipticity diagnostic are below 1 in the relevant parameter ranges. As the quotient in Eq. 6 is dominated by the error estimated on the data $\sigma_{d,ij}$, this behavior of the ellipticity diagnostics suggests issues in estimating the scatter of the dataset.

Finally, the magnitude-size plane is the most sensitive diagnostic to changes in the magnitude and size distribution parameters σ and θ of the galaxy population. Hence, the combined χ_{red}^2 -curve is mostly driven by this diagnostic.